You'll be working with the NYC Yellow Taxi Trip Dataset for January 2023. This dataset contains information about taxi rides including pickup/dropoff times, locations, fares, and tips.

https://www.nyc.gov/site/tlc/about/tlc-trip-record-data.page

Steps:

Download the January 2023 Yellow Taxi Trip data from the NYC TLC website

Create a PostgreSQL database named nyc\_taxi\_db

Design and create a table schema for the taxi data

Load the data into your PostgreSQL database

Load the taxi data into a pandas DataFrame and perform basic cleaning:

* Remove rows with missing values
* Remove outliers in trip distance and fare amount
* Convert datetime columns to proper datetime type

Create new features:

* Trip duration in minutes
* Speed (miles per hour)
* Time of day category (morning/afternoon/evening/night)
* Is weekend (True/False)

Create the requisite fields in the database and populate with the new features created.

Using matplotlib and seaborn, visualize the trends based on the time-series data:

* Plot daily total revenue
* Add a 7-day moving average trendline
* Annotate major peaks and troughs